DOCUMENT RESUME

ED 385 989 EC 304 191

AUTHOR	Thurlow, Martha
TITLE	Inclusion of Transition-Age Students with Disabilities in Large-Scale Assessments.
INSTITUTION	National Transition Network, Minneapolis, MN.
SPONS AGENCY	Office of Special Education and Rehabilitative Services (ED), Washington, DC.
PUB DATE	Jul 95
CONTRACT	H158G200002
NOTE	17p.
AVAILABLE FROM	National Transition Network, Institute of Community Integration (UAP), University of Minnesota, 86 Pleasant Steet, S.E., 430 Wulling Hall, Minneapolis, MN 55455 (available in alternative formats upon request).
PUB TYPE	Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Standards; *Disabilities; *Educational Assessment; *Educational Policy; Eligibility; Guidelines; High Schools; High School Students; Inclusive Schools; Mainstreaming; *Policy Formation; State Standards; Student Evaluation; *Student Participation; Testing Problems
IDENTIFIERS	*National Assessment of Educational Progress

ABSTRACT

This monograph describes the functions of large-scale educational assessments and examines existing and developing policies as they pertain to transition-age youth with disabilities. It stresses that the level of representation of students with disabilities in such assessments will have impact on their level of representation in national and state policy and service decisions. The first section looks at the use of large-scale assessments for system accountability and statewide decision making. Tables and figures list national data collection programs, show overall inclusion and exclusion rates, and indicate which states include students with disabilities. The second section addresses existing policy, with discussion of national policy (including the exclusion guideline used by the National Assessment of Educational Progress) and of state policy, noting the great variation in written guidelines on inclusion/exclusion guidelines. The third section, on policies in the making, identifies national activities such as the development of assumptions underlying inclusion guidelines, a survey to determine current inclusion/exclusion practices, and suggestions for revising National Assessment of Education Progress guidelines. State activities in developing policy are also addressed. A final section on issues for transition-age students stresses the need to clarify the role of the individualized education program and to focus on the goal of an individual student's education. (Contains 31 references.) (DB)

ED 385 989

National Transition Network

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Inclusion of Transition-Age Students with Disabilities in Large-Scale Assessments

July 1995

BEST COPY AVAILABLE

EC304191

Collaborating with states to enhance transition services and adult life outcomes for youth with disabilities

National Transition Network

Inclusion of Transition-Age Students with Disabilities in Large-Scale Assessments

July 1995

Collaborating with states to enhance transition services and adult life outcomes for youth with disabilities

This publication was prepared for the National Transition Network
by Dr. Martha Thurlow, Assistant Director of the
National Center on Educational Outcomes for
Students with Disabilities, at the University of Minnesota.

The National Transition Network is a collaboration of the University of Arkansas; Colorado State University; University of Illinois, Urbana-Champaign; University of Minnesota; PACER Center; Technical Assistance for Parent Programs, and the University of Vermont. Collaborators of the National Transition Network and the University of Minnesota are equal opportunity employers and educators.

Additional copies of this publication may be obtained by contacting:

National Transition Network
Institute on Community Integration (UAP)
University of Minnesota
86 Pleasant Street SE
430 Wulling Hall
Minneapolis, MN 55455
(612) 626-8200
FAX: (612) 626-7956

This publication is available in alternative formats upon request.

The development of this publication is supported by Cooperative Agreement H158G20002 from the Office of Special Education and Rehabilitative Services, U.S. Department of Education.

Table of Contents

Functions of Large-Scale Assessments ____ 1

System Accountability ____ 1

Assessments Used for Statewide Decision Making ____ 3

Assessments Used for Individual Decision Making ____ 4

Existing Policy ____ 5

National Policy ____ 5

State Policy ____ 6

Policies in the Making ____ 6

National Activities ____ 6

State Activities ____ 9

Issues for Transition-Age Students ____ 10

Conclusion ____ 11

References ____ 11

Inclusion of Transition-Age Students with Disabilities in Large-Scale Assessments

In less than a decade, there has been a dramatic increase in the amount of attention that our nation pays to assessments given "outside the classroom" (Office of Technology Assessment [OTA], 1992):

> *"Everyday millions of children take tests. Most are devised by teachers to see how well their pupils are learning and to signal to pupils what they should be studying. . . . It is another category of test, however—originating outside the classroom, usually with standardized rules for scoring and administration—that has garnered the most attention, discussion, and controversy (OTA, 1992, p. 1)."*

Assessment for instructional planning has, for the time being, taken the back burner in national and state discussions, while assessment for accountability has moved to the front. For the most part, large-scale assessments are used to monitor the educational system or determine certain educational opportunities.

Large-scale assessments are data-collection efforts in which large numbers of individuals are assessed. The assessment results are then used to describe educational status, make decisions about individuals, and may eventually result in the development or revision of existing state and national policies.

As educational reform gathers speed, and as greater emphasis is put on assessments for making decisions that may affect life opportunities, it is important to better understand the policies and issues that surround the inclusion of transition-age students with disabilities in large-scale assessments. The level of representation of students with disabilities in large-scale assessments will impact their level of representation in national and state policy and service decisions. This publication provides a brief description of the functions of large-scale assessments and examines existing and developing policy as they pertain to transition-age youth with disabilities.

Functions of Large-Scale Assessments

As the push for educational reform has increased, so has the perceived importance of large-scale assessments (ETS, 1994; OTA, 1992). Large-scale assessments are used at the national level, the state level, and sometimes the local level. The Office of Technology Assessment (OTA, 1992) described three purposes of assessment:

- to aid teachers and students in the conduct of classroom learning.
- to monitor systemwide educational outcomes.
- to make informed decisions about the selection, placement, and credentialling of individual students. (p. 8)

Usually, large-scale assessments focus on the latter two of these purposes. The Policy Information Center at ETS (1994), using data from the State Student Assessment Database (see Bond, 1994), organized the primary functions of state assessment programs into five purposes:

- Accountability.
- Instructional Improvement.
- Program Evaluation.
- Student Diagnosis.
- High School Graduation.

Large-scale assessments rarely are used for instructional planning/improvement or for diagnosing students. Most large-scale assessments focus on either system accountability *(describing educational status)* or selection, placement, and credentialling *(decision making)*. The two broad purposes of describing educational status and making decisions are the focus of much controversy, particularly in relation to educational policy.

System Accountability

System accountability drives much of the assessment that occurs nationally and in statewide assessments. The need for system accountability derives from the need for educational accountability and public information.

More than a dozen data collection programs of a recurring nature are conducted by the U.S. Department of Education, through the National Center for Education Statistics (NCES), and more are in the developmental stage (e.g., Early Childhood Longitudinal Study). Of these existing data collection programs, all include secondary level students or individuals who have recently left school as part of their sample, or as the total sample. Some

are longitudinal in nature, others are not. The same holds true for at least 15 data collection programs sponsored by agencies such as the Department of Labor, Department of Justice, Department of Health and Human Services, and the American Council of Education (McGrew et al., 1992). The names and sponsoring agencies of national data collection programs that gather information about secondary school students are presented in **Table 1.**

The National Assessment of Educational Progress (NAEP) is among the most well-known of the national education data collection programs. Known as our nation's report card, this assessment has changed since its inception in 1969 from an assessment without accountability pressures that provided information for the nation and regions of the U.S.(NAGB, 1992), to an assessment able to provide data linked to national and state policies, and considered to be a primary source of data for assessing the effects of educational reforms. NAEP describes the status of students in the United States, and on a trial basis, the performance of students in specific states on tests of achievement and related contextual variables (amount of time spent reading, watching TV, family configuration, etc.). NAEP is administered to students in grades 4, 8, and 12.

Although the NAEP data collection program started in the sixties, it has been given much greater emphasis in recent years, to the point where funds are being allocated to administer assessments more frequently than every other year (which was the previous schedule), and sometimes in more than one content area during the same year (Geenen, Shin, Thurlow, & Ysseldyke, 1994). NAEP is also the primary data set used to document our nation's progress toward the national education goals (National Education Goals Panel, 1991, 1992, 1993). Because the state NAEP data are being published in newspapers and other reports, the consequences ("stakes") associated with this assessment are perceived to have been raised considerably—they are considered by some to be high-stakes assessments.

Table 1

National Data Collection Programs that Collect Data on Transition-Age Students

Title of Data Collection Program	Sponsoring Agency
Advanced Placement Tests	The College Board
Beginning Postsecondary Students	Department of Educaton
Current Population Survey	Department of Commerce
General Education Development Testing	American Council of Education
High School and Beyond	Department of Education
Integrated Postsecondary Education Data System	Department of Education
International Assessment of Educational Progress	Department of Education
International Evaluation of Educational Achievement	Department of Education
Longitudinal Study of American Youth	National Science Foundation
Monitoring the Future	Dept of Health & Human Services
National Adolescent School Health Survey	Department of Health & Human Services
National Adult Literacy Survey	Department of Education
National Assessment of Educational Progress	Department of Education
National Crime Survey	Department of Justice
National Education Longitudinal Study	Department of Education
National Health and Nutrition Examination Survey	Department of Health & Human Services
National Health Interview	Department of Health & Human Services
National Household Education Survey	Department of Education
National Household Survey of Drug Abuse	Department of Health & Human Services
National Longitudinal Transition Study	Department of Education
Nat'l Survey of Personal Health Practices and Consequences	Department of Health & Human Services
National Survey of Family Growth	Department of Health & Human Services
Transcript Study	Department of Education
Workforce Participation Survey	Department of Labor
Workplace Literacy Assessment	Department of Labor
Young Adult Literacy Survey	Department of Education
Youth Risk Behavior Surveillance System	Department of Health & Human Services

October 1994

The increased emphasis on the NAEP data collection program has not, however, resulted in adequate data on students with disabilities. In fact, recent analyses estimate that nearly 50% of students with disabilities (McGrew, Thurlow, & Spiegel, 1993), or about 5% of all students (National Academy of Education, 1992, 1993) were not included in the 1990 and 1992 administrations of NAEP. Furthermore, it has been found that there is tremendous variability in the rates of exclusion of students with disabilities from state-level NAEP[1] administrations. The graph in **Figure 1** displays some of this variability. For instance, more than 60% of students with disabilities were excluded in nine states, almost as many states in which 40% or less were excluded. With this kind of variability, serious questions must be raised about the usefulness of state-level NAEP data (National Academy of Education, 1992, 1993). Furthermore, this variability raises serious questions about statewide assessments, since state practices in administering NAEP are likely to reflect state practices in their own assessments, with possibly even more exclusion from the state's own assessments because there is not an external agent saying, "when in doubt, include the student."

The only major national data collection program, of those listed in Table 1, to specifically focus on students with disabilities is the National Longitudinal Transition Study, which is sponsored by the U.S. Department of Education, Office of Special Education Programs. This data collection program followed 1987 base year students in special education for two waves of follow-up data collection in order to provide information on the transition of youth with disabilities from secondary school to early adulthood (see Wagner, D'Amico, Marder, Newman, & Blackorby, 1992; Wagner, Newman, D'Amico, Jay, Butler-Nalin, Marder, & Cox, 1991).

Thus, other than in selected special studies of students with disabilities, high exclusion rates and variable exclusion rates are significant problems for national data collection programs and significant barriers to having useful national descriptive data on students with disabilities.

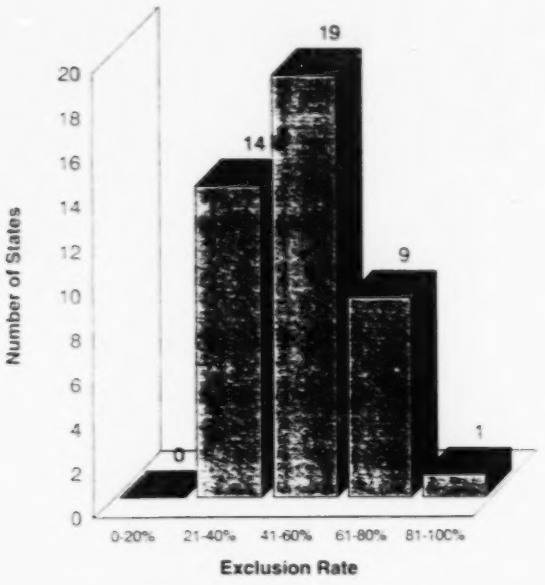

Figure 1

Number of States in Which Varying Percentages of Students with Disabilities Were Excluded from State-Level NAEP (1990)

NCEO

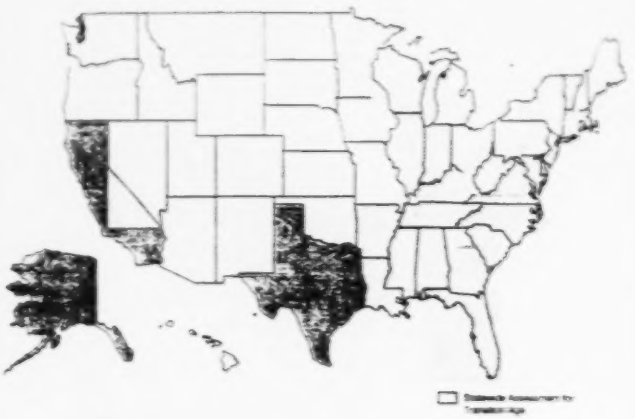

Figure 2

States with Statewide Assessments Used with Transition-Age Students for System Accountability

NCEO

Assessments Used For Statewide Decision Making

States also have used data collection programs for descriptive purposes. These data collection programs typically focus on academic achievement. They are administered differently than accountability assessments, both in terms of the nature of the assessment (multiple choice, portfolio, etc.) and the ways in which students are sampled. Some states collect data on "all" students in the public schools, some test students in a sample of districts, and some states collect data only from districts that volunteer to be part of the assessment. Some states administer assessments at all grades, some at only selected grades. **Figure 2** identifies those states that have statewide assessments used for descriptive purposes that include transition-age students.

State data are increasingly being used to provide the public with information and to set policy. *To what extent are students with disabilities included in these assessments?* This question has been more difficult to answer at the state level than at the national level. While there has been improvement over the past few years in the knowledge states have about the number of students with disabilities included, there are still relatively few states able to report the number of students with disabilities included in the statewide assessment (Shriner, Spande, & Thurlow, 1994). **Figure 3** is a summary of the estimated percentages of students with disabilities assessed in statewide assessments.

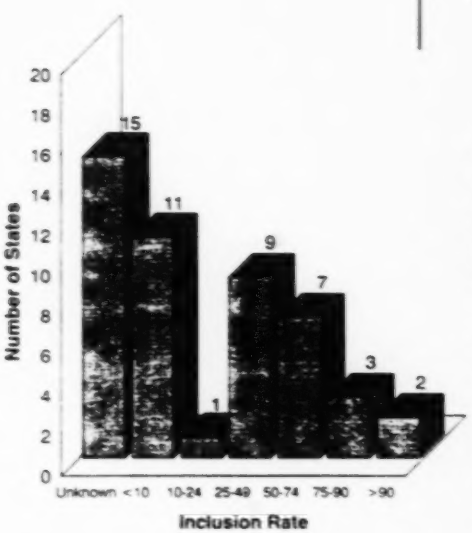

Figure 3

Number of States in Which Varying Percentages of Students with Disabilities were Assessed in Statewide Assessment NCEO

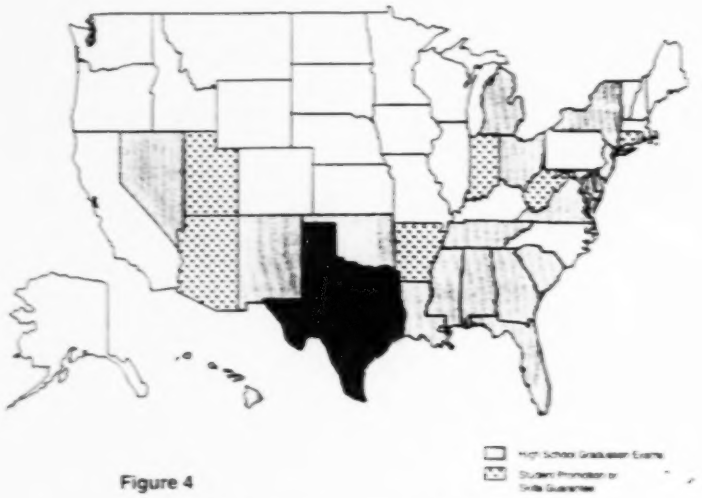

Figure 4

States with High State Assessments Used with Transition-Age Students for Individual Decisions NCEO

Assessments Used for Individual Decision Making

Decision making about individuals, based on assessment results, occurs primarily at the state level. Decisions may be made at the group level (for example, the level of funding for a school district or school might be based on student performance levels). Decisions may also be made about individuals (administrators, teachers, or students). Decisions about groups or school personnel are mentioned only briefly here because they affect only tangentially transition-age students with disabilities.

Decision making about individuals based on assessment results does not usually occur at the national level. National teacher certification exams such as the National Teacher Examination (NTE) have been developed, but are used only if a state decides to do so as part of its certification procedures (see McLaughlin, 1993 for information on this and other teacher certification standards and examinations). The national assessments that are used to make decisions about individual students are those in which students elect to participate (e.g., Scholastic Achievement Test [SAT], or American College Testing [ACT]).

Statewide assessments are often used to make decisions about individual students. These assessments are the "minimum competency tests" and "graduation exams" that students must pass in order to be awarded a high school diploma. Sometimes, exams are administered in earlier grades (usually grade 11) to make sure students are on the right track to pass the graduation exam. These are high- stakes assessments for students because they can determine whether the student receives a high school diploma, which in turn has a significant impact on life opportunities, an impact that needs to be evident in order to "trigger the due process protections of the fourteenth amendment in civil litigation" (Phillips, 1992). We currently are in the middle of a turnover from minimum competency testing (which was used in 40 states at one time; DeStefano & Metzer, 1991) to high school graduation or proficiency exams. **Figure 4** shows those states that currently have some version of high stakes assessment for transition age students.

The number of large-scale statewide assessments is likely to increase significantly in the next few years because the Goals 2000 education reform law signed

by President Clinton in March 1994 requires that states set high standards for all students, develop assessments for those standards, and then monitor progress on standards. It is very clear in this law that all students means *all* students, as reflected in the Act's definition:

> *. . . students or children from a broad range of backgrounds and circumstances, including disadvantaged students and children, students or children with diverse racial, ethnic, and cultural backgrounds, American Indians, Alaska Natives, Native Hawaiians, students or children with disabilities, students or children with limited-English proficiency, school-aged students or children who have dropped out of school, migratory students or children, and academically talented students and children.*

The law specifically prevents the use of these assessments for at least five years for making decisions about individual students, noting especially that the assessments should not be used to decide whether a student receives a diploma or proceeds to the next grade in school. After five years (approximately the amount of time from when a student is in seventh grade to when the student is in twelfth grade) the state may opt to use assessments to decide. When assessments are administered to make decisions about students, the question of inclusion of students with disabilities in the assessments becomes less clear-cut. Issues that arise when considering the inclusion of students with disabilities in these high stakes assessments is discussed in the section "Policies in the Making."

Existing Policy

National data collection programs and statewide assessments currently have policies in place regarding the inclusion of students with disabilities in assessments. Although many of these policies are now being questioned and revised, they provide a framework for examining the direction of policy change.

National Policy

In general, national data collection programs have included students with disabilities when the collection procedure has involved getting information from another individual (an informed respondent) or from existing records such as school files (McGrew et al., 1993). When the assessment involves administration of a test (usually an achievement test), some type of written guideline is used to help make judgments about whether an individual student is to be included in the assessment. The National Assessment of Educational Progress (NAEP), the most relevant recurring national education data collection program that includes transition-age students, has for several years used the guideline that a student on an Individualized Education Program (IEP) may be excluded if—

> *the student is mainstreamed less than 50 percent of the time in academic subjects and is judged to be incapable of taking part in the assessment or the IEP team has determined that the student is incapable of taking part meaningfully in the assessment.*

As noted previously, this guideline has resulted in extreme variability in inclusion rates. The variability reflects to some degree both vagueness in the guideline (e.g., what does it mean to "take part meaningfully?"—does it mean "perform well?") and inconsistencies in application of the guideline (e.g., it is easier to simply exclude all IEP students than to try to look at each student's schedule and to get the IEP team together to determine the student's capability for meaningful participation).

In a recent meeting with National Center for Education Statistics (NCES, March 1994) experts in assessment and/or disability issues from the National Center on Educational Outcomes (NCEO) summarized some of the primary reasons for the exclusion of students with disabilities (see Ysseldyke, Thurlow, McGrew, & Shriner, 1994; Ysseldyke, Thurlow, McGrew, & Vanderwood, 1994):

- **Vagueness** in assessment guidelines that leave the decision about student participation in the assessment up to a local decision-making process, which leads to differential implementation of the guidelines.
- **Variability** in monitoring the extent to which the intent of the guideline is followed.
- **Sampling** plans that systematically exclude students who are in separate schools and students who are not in graded programs.
- **Nonavailability** of accommodations in assessment materials and procedures.
- **Altruistic** motivations to lessen emotional distress to students who are not expected to do well.
- **Incentives** created by the desire to have a school or state look good in comparison to others in the state or nation.

These reasons apply to many other national (and state) assessments as well.

Evidence of the tendency to exclude students with disabilities was discovered and addressed competently in the National Education Longitudinal Study (Ingels, 1993). This national data collection program followed students starting in 8th grade in 1988 then picking them up again in grade 10 (or the grade they were in after two years—1990), and again in grade 12 (or where they were in 1992). It was discovered that students appeared to be excluded categorically, even though

guidelines had indicated this should not be done. At the first follow-up (1990), investigators went back to look at the eligibility of those students who had been excluded during the first year. It was found that 53% of previously excluded students did not meet criteria for exclusion, 40% still were ineligible and 7% could not be located. While many of the newly eligible students in 1990 were students with language exclusions originally, 39% of physical barrier exclusions and 42% of mental barrier exclusions were classified as eligible in 1990. These data led Ingels (1993) to "support the contention that a large number of students who could successfully have participated were excluded by their schools" (p. 11).

These and other examples have raised questions about the appropriateness of existing policy at the national level and its implementation at the state and local levels. The significant emphasis placed on such assessments examination of ways in which these policies can be altered, and hopefully, improved.

State Policy

While it is quite clear that national policy is in need of revision, both in terms of inclusion guidelines, and also in terms of allowing accommodations during assessments (NAEP has allowed no accommodations), needs for existing state policies are much less clear. This is due, to a large extent, to the extreme variability in both the nature of existing guidelines and the extent to which accommodations are used during these assessments.

An examination of the *written* guidelines of states on inclusion of students with disabilities in assessment confirms that states do vary considerably (Thurlow, Shriner, & Ysseldyke, 1994; Thurlow, Ysseldyke, & Silverstein, 1993). The five most frequent types of criteria that appear in these written guidelines are:

- Characteristics of the student's program or curriculum.
- IEP specification.
- Need for appropriate accommodations.
- Characteristics of the student.
- Parent or guardian opinion.

While not as frequent, written guidelines also included criteria that focused on the effect of the assessment on the student and the effect of including the student on the test results. What is not reflected in this listing is the fact that many states started with the NAEP guidelines when developing their policies about including students with disabilities in assessments. In general, the trend toward high exclusion rates was found in these states, just as it was for NAEP assessments.

State policies on statewide assessments also include the notion of providing for the use of accommodations by students with disabilities. The variability in the accommodations that are allowed is tremendous. Of particular interest is the finding that what is prohibited in one state may be recommended in another. For example, Tennessee specifically prohibits the use of more time as an accommodation, whereas eight other states recommend more time as an appropriate accommodation for certain students. Across states with written guidelines, the most frequently appearing as approved are:

- Braille version.
- Large print version.
- Individual setting.
- IEP defined.
- Small group setting.
- More time.

This list, as might be expected, is fluctuating as states change their guidelines on the use of accommodations in statewide assessments.

Policies in the Making

Clearly, policies regarding assessment are in transition. This is true in terms of the format of assessments (Shriner, Spande, & Thurlow, 1994; Thurlow, 1994), and the content of assessments, as well as in the consideration of the participation of students with disabilities.

National Activities

In recent months, NCES has started to look seriously at ways to include more students with disabilities in NAEP (Phelps, 1994). With evidence from NELS of the possibility of doing so successfully, and evidence from NAE and NCEO of the need to do so, national leaders have expressed their commitment to do so in responsible ways.

One of the first steps in moving toward increased inclusion of students with disabilities in assessment was to lay the issues out and discuss them. This was done at the March 1994 meeting of staff representing NCES, ETS (the contractor that develops and conducts NAEP), and NCEO, along with several experts in the fields of assessment, evaluation, and/or disability. First, the group agreed on several general assumptions that should underlie guidelines for inclusion in

NAEP and accommodation decisions. These assumptions were:

- Accuracy and fairness should characterize the assessment.
- Assessment should provide information on students with disabilities.
- Assessment procedures should be sensitive to the needs of students with disabilities.
- Assessment should make clear that the same high standards are expected of all students.
- Assessment should be characterized by practicality and cost effectiveness.
- Assessment should be consistent with students' instructional programs and accommodations.

It was also noted at this meeting that there is a need to be clear about the characteristics of students. There are 13 federal categories of students with disabilities, and within each category students demonstrate a wide range of skills and abilities. It is estimated that as many as 85% of the nearly 5 million students who are now considered eligible for special education services (i.e., they are on IEPs) could take large-scale tests without adaptations or accommodations. These students include many of the students with learning disabilities, emotional or behavioral disabilities, and some with mental retardation. One factor that limits the participation of some of these students in NAEP is the lack of items appropriate for low-functioning students. Of course, not all students with disabilities are low functioning.

Another major discussion point was the use of the assessment results. It was noted that individual scores on students and schools are not obtained in NAEP (and, while it is possible to derive district scores, doing so is still prohibited; NAGB, 1992). NAEP is designed to be used for descriptive purposes. There is a need to know how students with disabilities are doing in the nation's schools. Since NAEP is considered to be the nation's "report card," students with disabilities need to be included and information on their performance needs to be included in reports.

It was also noted at the March meeting that the current NAEP guidelines for making decisions is problematic in at least two ways: (1) its use of a percentage of time in the mainstream setting; and (2) its reliance on the "IEP team" (or some designated person) to make decisions about "meaningful participation" in assessments. Percentage of time in the mainstream is not a good indicator of a student's instructional program, level of skill development, or ability. There are too many other factors that enter into mainstreaming decisions. The IEP team provision allows too much slippage in the team decision-making process. Frequently, the IEP terminology is interpreted to mean that any student on an IEP should be excluded from testing. Sometimes decisions are made solely on the basis of the student's category of disability. Considerable training may be needed before IEP teams cound decide whether students should take tests.

Recommendations for revising the terminology in the NAEP guidelines addressed these two problematic aspects of the guidelines: **First**, rather than using a percentage of time measure, a better indicator would be correspondence between the content the test is intended to measure and the type of curriculum for the students. Students who are working toward outcomes other than those measured by the assessment (e.g., functional skills) should participate in a different assessment. The type of curriculum rather than the setting should be the factor that determines the content of assessment. **Second**, rather than referring to the IEP, it would be better to identify skills needed to take the assessment. School building administrators could be provided with a checklist of factors to consider in making inclusion/exclusion decisions.

In addition, specific recommendations were made with regard to students with disabilities for: (1) development of NAEP assessments, (2) administration of NAEP, and (3) monitoring participation in NAEP. The specific recommendations, summarized from Ysseldyke, Thurlow, McGrew, and Vanderwood (1994) are presented in **Table 2** (page 8).

Another important step in moving toward increased inclusion of students with disabilities in assessments was to gain a better understanding of exactly what was happening in the field. In spring 1994, as a follow-up to the administration of the NAEP grade 4 reading assessment, NAE asked several questions about students on IEPs who had been excluded from the assessment. Among the questions were:

- Who made the decision to exclude?
- What were the reasons for the decision?
- What kinds of modifications would have to be made for the students to participate in the assessment?

In addition to these questions, attempts were made to assess each excluded student, first using an assessment that tested a broad range of reading performance (starting with letter recognition), then (if a certain level was attained on the first assessment), using the NAEP reading assessment.

The results of this follow-up study have yet to be released. But it is expected that they will reinforce the notion that many more students with disabilities could and should be included in this national assessment than have been in the past.

NCES is moving ahead in its efforts to increase the inclusion of students with disabilities in NAEP. In field tests of the next NAEP to be administered, NCES plans to check different types of guidelines (for example, the use of a checklist or a

Table 2

Abridged Recommendations for NAEP from Ysseldyke, Thurlow, McGrew, and Vanderwood (1994)

Recommended Guidelines for Making Inclusion Decisions

1. **Include students with disabilities when trying out items** in order to identify problematic item formats and the need for more items at the lower end, for example.

2. **Include all students with disabilities in "taking" some form of the assessment**. When a sampling procedure is used for an assessment, the sample must be representative of **all** students.
   - **Allow partial participation in an assessment**
   - **Use [a different] assessment for some students**.

3. **Include students with disabilities in reporting of results**. Data on the performance of all students are needed. Therefore, scores must be reported for all students.

Recommended Guidelines for Making Accommodation Decisions

Not all students with disabilities will need accommodations during assessments. But modifications in assessments should be used when needed to increase the number of students with disabilities who can take tests. Accommodations and adaptations that teachers currently use with students during instruction and that are permitted by society should be used during assessments.

As new technologies and procedures for accommodations and adaptations are developed, they should be included in the array of possible accommodations and adaptations for instruction and testing.

Recommendations for Monitoring

Monitor adherence to the intent of the guidelines by making sure that no student is excluded who could participate if accommodations and adaptations were used. Do this by requiring a specific person in the district to sign off for each student who does not participate in the regular assessment and by having the student complete [a different] assessment or having someone provide information about the student. In addition:

- **Conduct follow-up studies of excluded students**
- **Remove incentives for exclusion**
- **Set up a panel to review requests for new forms of testing modifications**

October 1994

content-curriculum match criteria rather than percent time mainstreamed and meaningful participation).

State Activities

States continue to look at their guidelines related to both participation of students with disabilities in assessments and accommodations. During 1993 and 1994, approximately one-third of the states produced new guidelines on the inclusion of students with disabilities in statewide assessment programs and approximately one-third (not necessarily the same states) produced new guidelines on testing accommodations.

At the same time, states are rethinking the kinds of assessment they use. While many states have had statewide assessments to describe the general performance of students, many more are likely to be adding new descriptive assessments to respond to the requirements of the Goals 2000 legislation.

Statewide assessments that are high stakes in nature also are in great flux at this time. In the late 1980s, nearly 40 states had mandated some type of exit exam in high school (DeStefano & Metzer, 1991). Most of these were "minimum competency" assessments, designed to define "measurable standards of achievement expected of each student" (DeStefano & Metzer, 1991, p. 281). Today, these high stakes assessments are more often referred to as high school graduation tests, with the intent being to "raise the high school graduate's skills and knowledge to the higher level expected for success in a complex, demanding society and workplace" (Bond, 1993, p. 9).

In a discussion of issues and recommendations regarding high school graduation tests (Mehrens, 1993), the following recommendation was made in consideration of special education and limited English proficiency:

> ***Recommendation 33:*** *Enact an administrative rule regarding testing issues related to special education students and students with limited english (sic) proficiency.* (p. 41)

Furthermore, it was stated that:

> *The state board of education needs to decide what to do with special education students. This decision includes the possibility of exempting such students and providing special administrative procedures and adapted versions of the test for certain handicapping conditions* (see The Americans with Disabilities Act, 1990). (p. 41)

These suggestions are fairly vague. It is little wonder that states continue to be in a state of flux regarding statewide assessments and students with disabilities.

In May 1994, NCEO staff pulled in several state assessment and state special education directors to begin to discuss guidelines for the participation of students with disabilities and the use of accommodations in their statewide assessments. The assumptions were:

- Any time data are collected for the purpose of making policy or accountability decisions, include all students. Not all students need to take the same test.
- The critical question to ask when considering the use of a different assessment is why the student is in a different curriculum. Inclusion in the curriculum is the first critical decision that is made for a student as an IEP is developed. If the student is not in the regular curriculum, it is important to ask why not. Then questions about the assessment can be asked.
- State assessment programs are conducted for multiple purposes. There is a need to differentiate participation and accommodation decisions as a function of purpose.
- Accuracy and fairness should characterize state assessment programs.
- Assessment procedures should be sensitive to the needs of students with disabilities.
- Accommodations are used for equity, not advantage. Students who use accommodations during an assessment do so to be able to take an assessment on an equal playing field as other students who do not need accommodations. Accommodations are not provided to help the student with a disability do better than other students.
- Assessment programs should make clear that the same high standards are expected of all students. State advisory boards should decide the range of performance permitted for each content standard.
- Assessment should be characterized by practicality and cost effectiveness.
- Assessment should be consistent with students' instructional programs and accommodations.

The primary recommendations that were generated from this meeting were very similar to those generated at the meeting with NAEP staff (see Table 2). Yet, an important caveat appeared in the report on this meeting (Ysseldyke, Thurlow, McGrew, & Shriner, 1994):

> *Guidelines for making decisions about inclusion and accommodations could vary as a function of the impact of the assessment on the student. The changes in guidelines at the national level, for example, do not have a direct impact on the student and are therefore considered to be "low stakes." As states turn assessments into*

> *high stakes for a school or district or for the personnel in them, the motivation increases to exclude those students who are perceived to bring average scores down.*
>
> *When a statewide assessment is "high stakes" for the student, such as graduation exams are, then it is imperative that consideration be given to guidelines. This does not mean that students with disabilities should be excluded from the assessment, but rather that appropriate accommodations must be made.* (pp. 11-12)

These recommendations are not much more specific than those provided earlier by the North Central Regional Education Laboratory.

Issues for Transition-Age Students

Now is the time for policymakers, educators, and parents to unite on important issues like educational accountability and shared responsibility for students and the results of education. Separate assessment systems or systems that allow exclusions from participation in accountability systems reinforce the notion that *all* educators are not responsible for *all* students. It is important to address the purpose of an assessment when thinking about the issues that accompany the inclusion of transition-age students with disabilities in large-scale assessments.

It is critical to clarify the role of the IEP in making decisions about assessment (DeStefano, 1993). Some have suggested that every decision related to assessment should be made by the IEP team. Others have argued that the IEP, as currently implemented, is not valid for program monitoring or for student assessment (Olsen & Massanari, 1991).

When the purpose of a large-scale assessment is to *describe* the status of students in the educational system, why would it make sense for some students to be excluded? It is too easy for someone to say that the assessment experience is too stressful for the student. The stress and anxiety associated with assessment is related to the preparation and test taking training that are provided to the student, although even the best-performing students will admit to considerable anxiety and stress associated with some assessment situations. Should not all students have the benefit of preparation for assessment-like experiences? They are a part of the life experiences of all individuals. Isn't it "too much process" to require that a team of individuals familiar with a student make the decision about whether a student should be assessed in an accountability system?

When the purpose of assessment is to *make a decision* about an individual student, then several other critical issues must come into consideration. It is particularly important in this case to look for unintended consequences of participation or nonparticipation in an assessment. For example, some states have enacted the policy that any special education student given an exemption from a graduation exam receives a graduation diploma. As might be expected, these states are much more likely to experience increased parent requests for special education services for their transition-age child. Another unintended consequence of an exclusion policy that focuses on the content of the student's curriculum is that students may be shuffled into another curriculum so they will not be included in reports of assessment results or in counts of the number of students eligible to receive a high school diploma who actually did. These and other considerations are the foundation of the issues that arise.

Perhaps the first issue to address when considering a student's participation is whether there is agreement among the key stakeholders (including the student's parents and the student) about the goal of a student's education. If the goal is for the student to follow the same general course of instruction and to reach the same outcomes as other students, then it follows that the student should be required to demonstrate attainment of the same goals (such as is required in a high school exam). It may be that the student will need to demonstrate the goals in different ways, with some kind of accommodation in procedures, but meet the same basic requirements.

Now is the time to struggle with the issues that emerge as we move toward better educational accountability and the graduation of students capable of successful work experiences or post-secondary training. This struggle is particularly challenging as we consider the involvement of transition-age students with disabilities. We now know that it is not advantageous for most of our students to be segregated in their educational experiences from those of their peers without disabilities. These peers are the individuals we want to have interact with individuals with disabilities in future life communities and workplaces. Therefore, it is imperative that students with disabilities be included in the accountability systems for these students.

What does the future hold? We know that the emphasis on higher standards and the cry for students to "earn" their diplomas is going to continue. Increasingly, concerns are being voiced about the worthlessness of traditional high school diplomas (National Center on Education and the Economy, 1994b). The trend is for states to use certificates of initial mastery. For example, in Indiana students will soon be required to earn a "gateway certificate" before they can earn their high school diploma (National Center on Education and the Economy, 1994a). The gateway certificate is earned by passing an exam on knowledge in core subjects. Six other states are in the process of implementing similar programs.

How will students with disabilities fit into these plans? It is time to insist on being included in educational reform and accountability systems, and it is time to

decide how to best meet the needs of individual students without depriving them of the experiences needed to become functioning members of society.

Conclusion

What we have now is a hodgepodge of approaches to the inclusion of transition-age students with disabilities in large-scale assessments. This situation of mixed policies and practices has occurred for a variety of reasons. Regardless of these, it is now time for educators to rethink what is happening in their schools, districts, states, and the nation.

Think about your own state. *Do you know what its policies are on the participation of students with disabilities in large-scale assessments? What are its policies on the use of accommodations or adaptations in assessments? What are the policies on reporting student performance? Do practices in you state match the policies?*

If your state is ready to take a fresh look at the issues surrounding statewide assessments and other accountability mechanisms, there are five primary issues that need to be addressed when considering the inclusion of students with disabilities in large-scale assessments:

- Given the purpose of the assessment, is it important to include students with disabilities in it?
- What factors should be used to make decisions about including students with disabilities in large-scale assessments?
- What is the role of the IEP in making decisions about inclusion in assessments?
- What accommodations should be used during these assessments?
- What is the role of the IEP in making decisions about accommodations during assessments?

If your answers to these questions point to the need for an alteration in policies and practices, it is important to implement an approach to change that will limit the impact of costs associated with the change. And, it will be important to involve both the director or assessment and the state special education director in this discussion, along with other stakeholders in the results of the assessments.

References

Bond, L. (1993). Executive summary. In W. A. Mehrens, *Issues and recommendations regarding implementation of high school graduation tests*. Oak Brook, IL: North Central Regional Educational Laboratory.

Bond, L. (1994). *Surveying the landscape of state educational assessment programs*. Washington, DC: Council for Educational Development and Research.

DeStefano, L. (1993). *The effects of standards and assessment on students in special education*. (Synthesis Report 10). Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

DeStefano, L., & Metzer, D. (1991). High stakes testing and students with handicaps: An analysis of issues and policies. *Advances in Program Evaluation, 1A*, 281-302.

ETS. (1994). *Testing in America's schools* . Princeton, NJ: Policy Information Center, Educational Testing Service.

Geenen, K., Shin, H., Thurlow, M., & Ysseldyke, J. (1994). *Synthesis report update 1993: Recent activities in national goals, standards, and tests* (Synthesis Report 8). Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Ingels, S. (1993). *Strategies for including all students in national and state assessments: Lessons learned from a national longitudinal study*. Paper presented at the National Conference on Large Scale Assessment, Albuquerque.

McGrew, K. S., Thurlow, M. L., Shriner, J. G., & Spiegel, A. N. (1992). *Inclusion of students with disabilities in national and state data collection programs* (Technical Report 2). Minneapolis, MN: National Center on Educational Out comes, University of Minnesota.

McGrew, K. S., Thurlow, M. L., & Spiegel, A. N. (1993). An investigation of the exclusion of students with disabilities in national data collection programs. *Educational Evaluation and Policy Analysis, 15* (3), 339-352.

McLaughlin, M. J. (1993). *Teachers and educational restructuring*. Paper developed for the Goals 2000 Special Education Work Panel.

Mehrens, W. A. (1993). *Issues and recommendations regarding implementation of high school graduation tests*. Oak Brook, IL: North Central Regional Educational Laboratory.

NAGB. (1992). *Discussion paper on the future of the National Assessment of Educational Progress*. Washington, DC: National Assessment Governing Board, Ad Hoc Committee on the Future of NAEP.

National Academy of Education (1992). *Assessing student achievement in the states* (First report of the National Academy of Education Panel on the Evaluation of the NAEP Trial State Assessment). Stanford, CA: National Academy of Education, Stanford University.

National Academy of Education (1993). *The trial state assessment: Prospects and realities* (Third report of the National Academy of Education Panel on the Evaluation of the NAEP Trial State Assessment). Stanford, CA: National Academy of Education, Stanford University.

National Center on Education and the Economy. (1994a). *States begin developing the certificate of initial mastery*. Rochester, NY: NCEE.

National Center on Education and the Economy. (1994b). *The certificate of initial mastery : A primer.* Rochester, NY: NCEE.

National Education Goals Panel. (1991). *The national education goals report 1991*. Washington, DC: U.S. Government Printing Office.

National Education Goals Panel. (1992). *The national education goals report 1991*. Washington, DC: U.S. Government Printing Office.

National Education Goals Panel. (1993). *The national education goals report 1991*. Washington, DC: U.S. Government Printing Office.

Olsen, K., & Massanari, C. (1991). *Special education program evaluation: What should states consider?* Lexington, KY: University of Kentucky, Mid-South Regional Resource Center.

OTA. (1992). *Testing in American schools: Asking the right questions*. Washington, DC: Congress of the United States, Office of Technology Assessment.

Phelps, M. (1994, May 16). ED officials press case for greater NAEP inclusion. Education Daily, 27 (93), 1, 3.

Phillips, S. E. (1992). *Testing condition accommodations for handicapped students*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Shriner, J. G., Spande, G. E., & Thurlow, M. L. (1994). *State special education outcomes 1993*. Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Thurlow, M. L. (1994). *National and state perspectives on performance assessment with students with disabilities*. Reston, VA: Council for Exceptional Children.

Thurlow, M. L., Shriner, J. G., & Ysseldyke, J. E. (1994). *Students with disabilities in the context of educational reform based on statewide educational assessments*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Thurlow, M. L., Ysseldyke, J. E., & Silverstein, B. (1993). *Testing accommodations for students with disabilities: A review of the literature* (Synthesis Report 4). Minneapolis, MN: National Center for Educational Outcomes, University of Minnesota.

Wagner, M., D'Amico, R., Marder, C., Newman, L., & Blackorby, J. (1992). *What happens next? Trends in postschool outcomes of youth with disabilities: The second comprehensive report from the National Longitudinal Transition Study of Special Education Students*. Menlo Park, CA: SRI International.

Wagner, M., Newman, L., D'Amico, R., Jay, E. D., Butler-Nalin, P., Marder, C., & Cox, R. (1991). *Youth with disabilities: How are they doing? The first comprehensive report from the National Longitudinal Transition Study of Special Education Students*. Menlo Park, CA: SRI International.

Ysseldyke, J. E., Thurlow, M. L., McGrew, K. S., & Shriner, J. G. (1994). *Recommendations for making decisions about the participation of students with disabilities in statewide assessment programs* (Synthesis Report 15). Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Ysseldyke, J. E., Thurlow, M. L., McGrew, K. S., & Vanderwood M. (1994). *Making decisions about the inclusion of students with disabilities in large-scale assessments*. (Synthesis Report 13). Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

Footnote

[1] Only since 1990 has NAEP been administered in a way that allows the reporting of data at the state level. Whether this should be done remains an issue, and further studies are being conducted.